J. mar. biol. Ass. U.K. (1974) 54, 559–563
Printed in Great Britain

COMPARING SPECIES DIVERSITY AND EVENNESS INDICES

BY C. HEIP AND P. ENGELS
Department of Zoology, State University of Ghent, Belgium

In a low diversity brackish water habitat the diversity of the copepod community is best measured with the Shannon–Wiener information function and its evenness by an index proposed by the first author $(e^H - 1)/(S - 1)$. This was shown by comparing the statistical behaviour of the more important diversity and evenness indices currently used.

INTRODUCTION

The current interest in diversity as one of the major parameters describing a community has led to an increasing number of published results and theoretical discussions during the last years. Diversity has variously been related to other attributes of the community or properties of the environment, among which time, spatial heterogeneity, stability, primary production, productivity, competition, predation, niche structure and evolution. Although this rather overwhelming amount of possible relations seems somehow too much of a good thing, the importance of diversity remains well established in current ecological theory. One of the most important applications of diversity indices is their usage in the biological assessment of pollution. This immediately raises the problem of comparing diversity indices within habitats in time or between habitats in communities. These comparisons are frequently made, yet the statistical significance of the observed differences or similarities is seldom mentioned. A test for these differences was formulated by Hutcheson (1970) but it apparently failed to attract the attention it deserves. In this study it was our purpose to investigate the distributions of the more important diversity indices and their evenness components in a low-diversity community of meiobenthic copepods of a shallow brackish water pond in northern Belgium. In this habitat the number of species is rather low, which greatly simplifies the analysis, and we hope that the statistical behaviour of the indices we find will hold for other, especially other low diversity communities as in polluted environments, as well. We chose copepods because they are taxonomically well known, the benthos because it is approximately two-dimensional and a shallow environment because of the possibility of accurate sampling.

MATERIAL AND METHOD

A wooden frame was dropped on the bottom where depth was about 7 cm. In the frame strings were stretched parallel to each other every 10 cm from one side to the other, in both directions; in this way a grid was formed with cells of 100 cm^2. A sample was taken in one of the corners everywhere two strings met; this was done with a core covering a surface area of 6 cm^2. The samples were fixed with 4 % formalin, brought to and elutriated in the laboratory. The copepods were extracted from the detritus under the dissecting microscope, determined and counted. It has been proven that this method has an efficiency close to 100 % (Heip, 1973, Thesis, Rijksuniversiteit Gent).

64 samples, forming a square of 80 cm on a side, were treated in this way. For each sample the diversity and evenness indices shown in Table 1 were calculated and the frequency distributions obtained where characterized by their mean ($\bar{x}$), standard deviation (s) and coefficient of variation ($s/\bar{x}$), and tested for normality by calculating skewness (g_1) and kurtosis (g_2).

RESULTS

The mean number of copepods was 119·55 per sample, which means that 7651 copepods were determined. The mean number of species was 4·09, total number being 6. The values of the most important diversity and evenness indices were calculated according to the formulae of Table 1; from this table several conclusions on the suitability of these indices can be drawn.

TABLE 1. DIVERSITY AND EVENNESS INDICES

(Mean, standard deviation, skewness and kurtosis.)

Diversity index	Formula	Total	$\bar{x}$	s	$s/\bar{x}$	g_1	g_2
Margalev	$Ma = (S-1)/\ln N$	0·56	0·66	0·19	29·6	0·09	3·18**
Simpson (a)	$1-SI = 1-\frac{\Sigma n_i(n_i-1)}{N(N-1)}$	0·16	0·16	0·07	42·5	0·39	3·85**
Simpson (b)	$1/SI = \frac{N(N-1)}{\Sigma n_i(n_i-1)}$	1·19	1·20	0·11	8·9	1·20**	6·16**
McIntosh	$MI = 1-\frac{\sqrt{\Sigma n_i^2}}{N}$	0·084	0·085	0·038	44·3	0·57	4·05**
Shannon–Weiner	$H = -\Sigma\frac{n_i}{N}\log_2\frac{n_i}{N}$	0·58	0·54	0·20	37·1	−0·09	2·93**
Brillouin	$B = \frac{N_1}{N}\ln\frac{N!}{n_1!n_2!\ldots n_S!}$	0·41	0·34	0·13	38·8	−0·11	2·93**
Number of species,	S	6·00	4·09	0·92	22·5	−0·19	2·84**
Evenness Index							
Simpson	$E = \frac{(1-SI)}{(1-SI_{max})}$	0·1941	0·2155	0·0896	41·9	0·83**	5·62**
McIntosh	$E = \frac{\sqrt{\Sigma n_i^2}}{\sqrt{((N-S+1)^2+S-1)}}$	0·1462	0·1695	0·0749	44·2	1·33**	7·67**
Lloyd & Ghelardi	$E = \frac{S}{S'}$	—	0·4213	0·0896	20·6	1·39**	6·66**
Pielou	$E = \frac{H}{H_{max}} = \frac{H}{\log_2 S}$	0·2258	0·2651	0·0922	34·8	0·73*	5·72**
Scaled	$E = \frac{H-H_{min}}{H_{max}-H_{min}}$	0·2230	0·1746	0·0956	54·8	0·63*	4·55**
Sheldon	$E = \frac{e^H}{S}$	0·2498	0·3709	0·0773	20·8	1·21**	5·07**
New index (Heip, 1974)	$E = \frac{e^H-1}{S-1}$	0·0998	0·1515	0·0664	43·7	0·07	0·28

N = total number of individuals.
n_i = number of individuals of the ith species.
S = number of species.
S' = number of species according to MacArthur's broken stick model yielding the observed diversity.
* Significant at 5 % level.
** Significant at 1 % level.
Total: value for all samples lumped together.

Diversity

All species diversity indices show a significant degree of positive kurtosis, but no skewness (except for $SI(b)$).

The variability of the species diversity indices is extremely low for $SI(b)$. It is between 20 and 30% for the number of species S and Margalev's index Ma, both indices which do not take the distribution of individuals over species into account. It is between 30 and 40% for H and B and between 40 and 50% for MI and $SI(a)$.

The conformity between the mean of the distribution and the value calculated for the total of all samples is best for $SI(a)$ and $SI(b)$ and for MI, worst for Ma and S. The last is obviously the consequence of the fact that the number of species will continually increase with sample size.

The close resemblance between SI and MI becomes obvious when noticing that $MI = 1 - \sqrt{SI}$, as can be shown easily. This is important with regard to Hill's (1973) unifying notation where diversity numbers N_a are defined as reciprocals of the $(a-1)$th root of a weighted mean of the $(a-1)$th powers of the proportional abundances of the S species. Hill (1973) shows that his diversity numbers of the 0th, 1st and 2nd order coincide with three important diversity measures, $N_0 = S$, $N_1 = e^H$ and $N_2 = 1/SI$. As MI is related to SI, it seems preferable to discard the MI-index in the future, as SI is the conceptually better index.

The coefficient of variation of H and B is about the same, but the conformity to the total sample is much higher for H. The values of these indices when calculated on a common (ln) base are nearly the same for the total sample ($H = 0{\cdot}405$ and $B = 0{\cdot}407$) but more different for the sample mean ($H = 0{\cdot}375$ and $B = 0{\cdot}336$).

In general, the better the variability of the index, the poorer its conformity with the total population or its statistical performance with regard to normality. Therefore, a compromise seems to lay in the middle two indices, H and B, with a variability around 38% and of these two H seems to be the better one with regard to conformity with the total sample. As H is the most widely used diversity index, this seems to be a rather lucky result. Its standard deviation as calculated from Hutcheson's (1970) method is 0·211, slightly higher than the observed 0·201 for our sample. This confirms Hutcheson's own observation that his formula yields a conservative test.

Evenness

All evenness indices show a considerable amount of both skewness and kurtosis, except for the index proposed by the first author (Heip, 1974). The first two indices coincide with their respective diversity indices and can only be used in connexion with those. The third index is a special case, but as MacArthur himself denounced his broken stick model, there seems to be little merit in continuing its usage. The last four indices are connected with the Shannon–Wiener information function H. Theoretically, Sheldon's index should be best as it results from dividing two of Hill's diversity numbers, indicating that evenness does not change when multiplying the number of individuals of all species with a constant and adding no species. Intuitively, this seems to be a necessary prerequisite for an evenness index (and, contrary to Hill's (1973) statement, the most widely used evenness index of Pielou shows this property). Sheldon's index shows this property also and should stabilize to a true community value when the size of the sample increases (Hill, 1973). However, as Heip (1974) pointed out, in low diversity communities were $H \to 0$, this index will increase instead of decrease because in most situations a decrease of H will be accompanied by a decrease in S. Because, when $H \to 0$, the lower limit of Sheldon's index is set by $1/S$, this limit will actually increase when $S \to 1$. This is the main reason for its rather small variability in the investigated habitat with its low number of species. The index proposed by Heip (1974) was derived from Sheldon's index keeping this in mind; it will vary between 0 when $H \to 0$, and 1 when $H = \ln S$. The statistical behaviour of this index eventually turned out to be far better than that of all other evenness indices, it being the only one showing no significant departure from normality.

Hill's (1973) supposition that a relationship of the form $N_1 = N_2 + a$ exists between diversity numbers of different orders and that therefore the difference $N_1 - N_2$ could be more characteristic of the community than its evenness N_2/N_1, does not hold in our case. The regression between $N_2 = 1/SI$ and $N_1 = e^H$ is $e^H = 1{\cdot}84\,(1/SI) - 0{\cdot}75$ ($r = 0{\cdot}965$).

DISCUSSION

We may conclude that in the investigated community diversity is best measured using the Shannon–Wiener information function H or e^H, which has $s = 0{\cdot}37\bar{x}$ and a normal distribution; evenness is best measured with the index proposed by Heip (1974) $(e^H - 1)/(S - 1)$, which has $s = 0{\cdot}44\bar{x}$ and a normal distribution. Comparison then becomes possible using these values of the standard deviation, or, in the case of H, using Hutcheson's (1970) formula for the variance of H:

$$\mathrm{Var}\,H = \frac{\Sigma p_i \ln^2 p_i - (\Sigma p_i \ln p_i)^2}{n} + \frac{S-1}{2n^2} + \frac{(-1 + \Sigma p_i^2 - \Sigma p_i^{-1} \ln p_i + \Sigma p_i^{-1}\,\Sigma p_i \ln p_i)}{6n^3} + \ldots$$

In extending these results to other communities we must be careful, as the results may be a function of the taxa examined, of the environment or the total number of species in the community. These questions can only be answered by survey of other environments and other kinds of organisms. Furthermore, patchiness can seriously affect the estimates obtained; in the investigated community there is patchiness for all species (Heip, 1974) but only to a slight degree. This patchiness seems not to be strong enough in this situation, but it is easy to imagine that it would make the use of diversity indices altogether impossible. Keeping all these restrictions in mind, we can only hope that these results are applicable to other communities in other environments as well, especially in other low diversity environments. A further test on the applicability of the Shannon–Wiener information function in the investigated community was done by investigation of the distribution of the mean of all 16 squares of samples with 30 cm on a side. The mean of these 16 values was 0·54, the standard deviation 0·09 and the distribution of the means showed no significant departure from normality ($g_1 = 0{\cdot}65$; $g_2 = 2{\cdot}17$). Five series of five samples drawn at random from the grid (with the aid of a table of random numbers) showed a distribution with mean 0·57, deviation 0·10 and no significant departure from normality ($g_1 = -0{\cdot}07$; $g_2 = 1{\cdot}75$). Both the random samples and the squares yield a mean which is drawn from a normally distributed population.

The first author acknowledges a grant as Aspirant of the Belgian National Foundation for Scientific Research (N.F.W.O.).

REFERENCES

HEIP, C., 1974. A new index measuring evenness. *Journal of the Marine Biological Association*, **54**, 555–7.

HILL, M. O., 1973. Diversity and Evenness: a Unifying Notation and its Consequences. *Ecology* **54**, 427–32.

HUTCHESON, K., 1970. A Test for Comparing Diversities based on the Shannon Formula. *Journal of Theoretical Biology* **29**, 151–4.